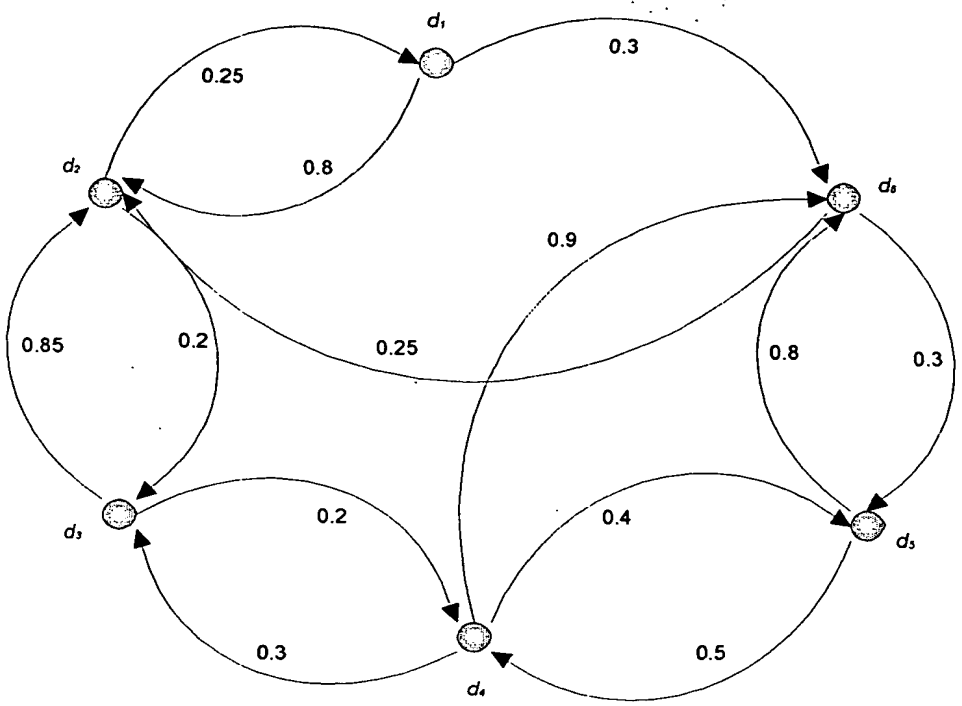


Figure 1

SCANNED, #14

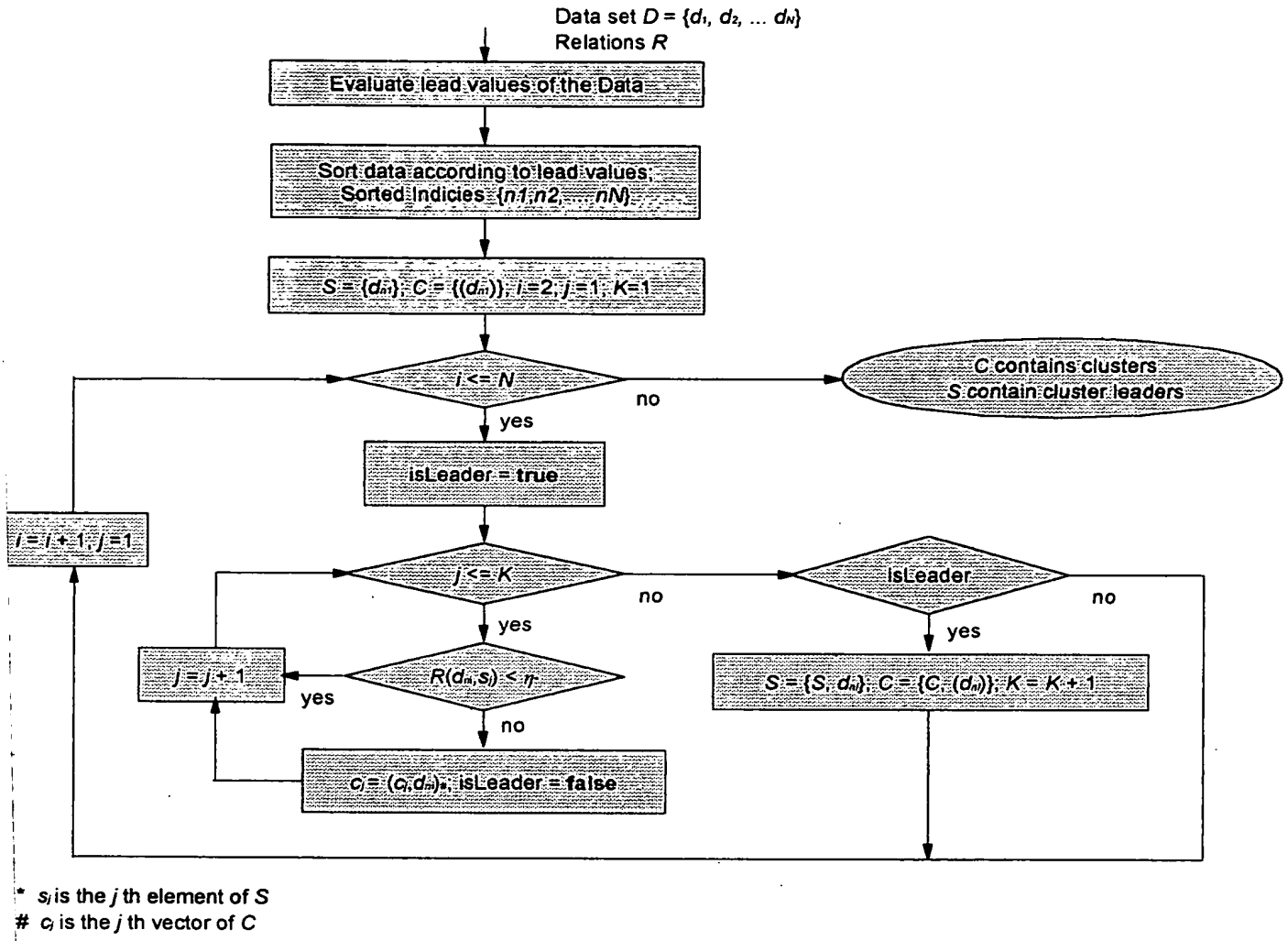

Data set D = {d1, d2, ... dN}
Relations R
Evaluate lead values of the Data
Sort data according to lead values;
Sorted Indicies {n1,n2,....nN}
S = {dn1}; C = {(dn1)}; i=2, j=1, K=1
i <= N
no
C contains clusters
S contain cluster leaders
yes
isLeader = true
i = i + 1; j =1
j <= K
no
isLeader
no
yes
yes
j = j + 1
R(dni, sj) < η
yes
S = {S, dni}; C = {C, (dni)}; K = K + 1
no
ci = (ci, dni)*; isLeader = false
* sj is the j th element of S
cj is the j th vector of C

Figure 2

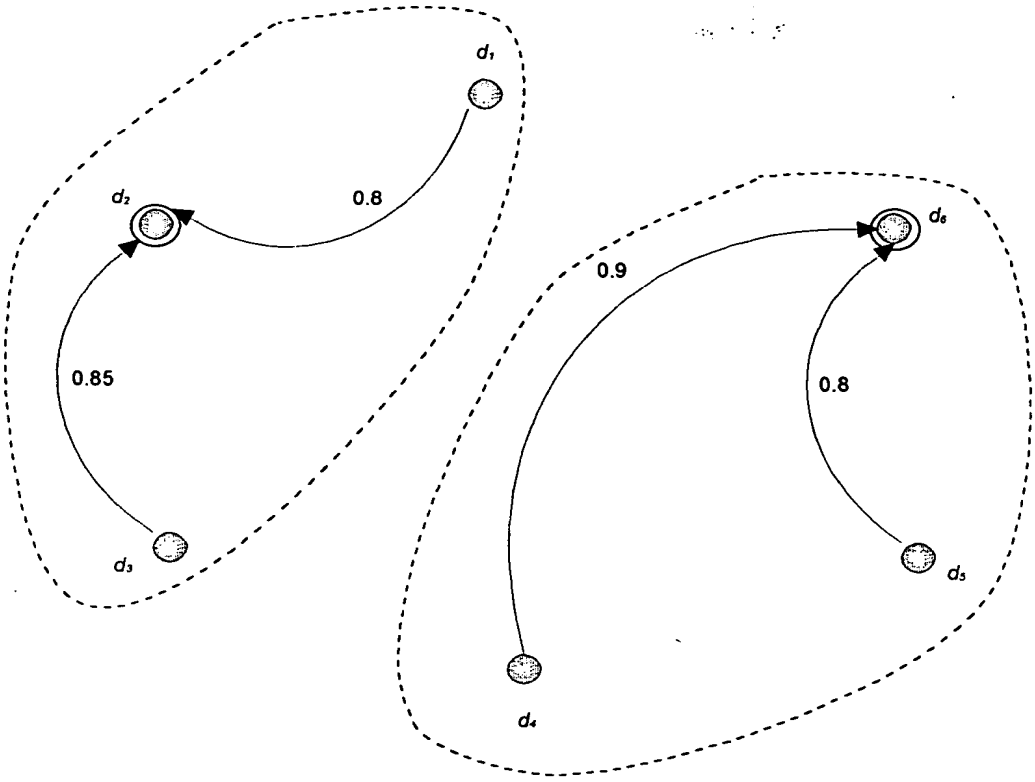


Figure 3

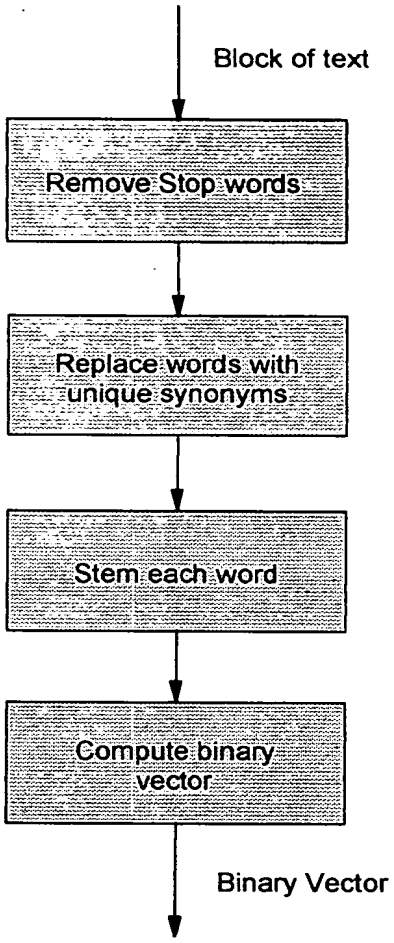

Block of text
Remove Stop words
Replace words with unique synonyms
Stem each word
Compute binary vector
Binary Vector

Figure 4

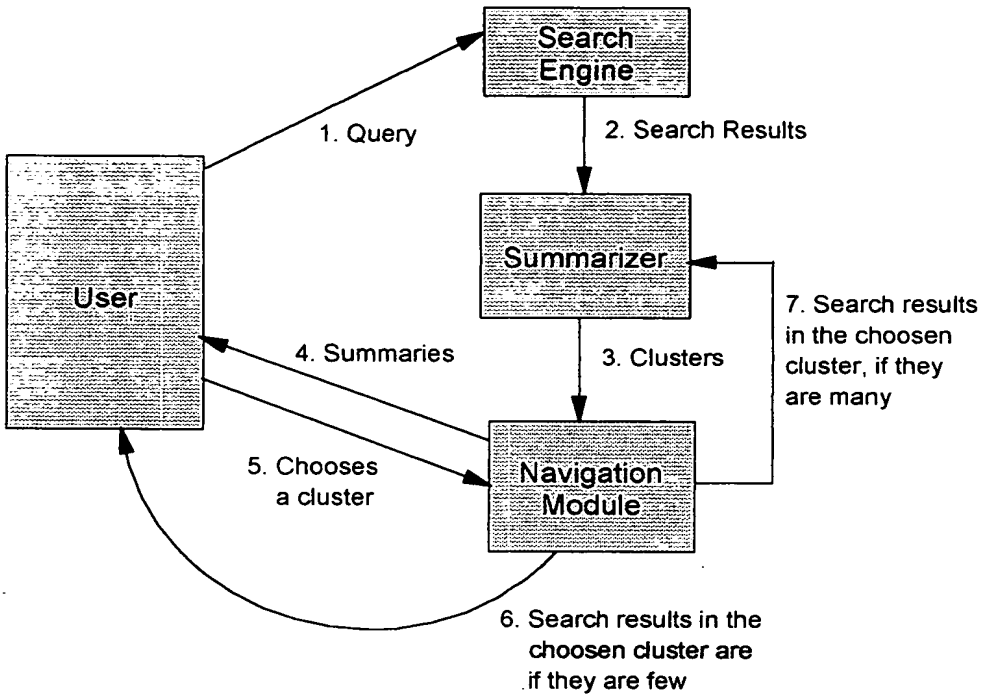

Search Engine
1. Query
2. Search Results
Summarizer
User
4. Summaries
3. Clusters
7. Search results in the choosen cluster, if they are many
5. Chooses a cluster
Navigation Module
6. Search results in the choosen cluster are if they are few

Figure 5

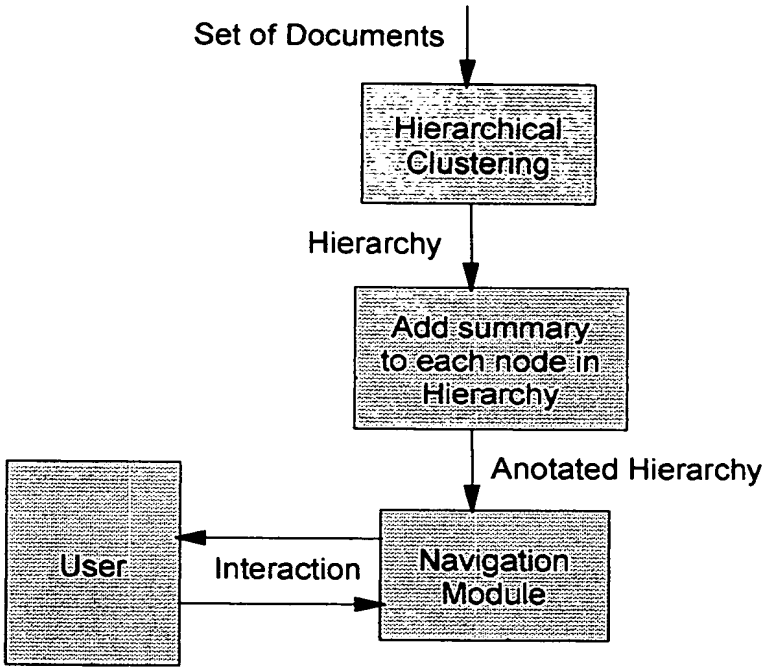

Set of Documents
Hierarchical Clustering
Hierarchy
Add summary to each node in Hierarchy
Anotated Hierarchy
User
Interaction
Navigation Module

Figure 6